

XAVIER AGENJO BULLÓN  
FRANCISCA HERNÁNDEZ CARRASCAL

# LA DIGITALIZACION DE LA BIBLIOTECA DE MENÉNDEZ PELAYO: I<sup>a</sup> FASE. MENÉNDEZ PELAYO Y EL "PROYECTO POLÍGRAFOS"

## LA DIGITALIZACIÓN DE LA BIBLIOTECA DE MENÉNDEZ PELAYO: Iª FASE. MENÉNDEZ PELAYO Y EL *PROYECTO POLÍGRAFOS*

Uno de los proyectos más importantes en la nueva fase de la Biblioteca de Menéndez Pelayo, tras su entrada en Internet a través de su página *web*, ha consistido no tanto en la conversión de sus catálogos como en una apuesta por la digitalización de sus fondos. Esta decisión tiene un especial significado ya que con ella se abre una línea de difusión de la información centrada en los documentos más que en el propio catálogo de la biblioteca. Con ello, siguiendo una de las corrientes más activas de las disciplinas humanistas, la biblioteca, en tanto que centro y foco de información, adopta un papel activo en la edición electrónica de textos, colaborando así con otros organismos científicos.

La primera fase del proyecto ha consistido justamente en la digitalización de las obras de Menéndez Pelayo, para lo cual se ha partido tanto de la edición de las *Obras completas* que se realizó conjuntamente con el CSIC<sup>1</sup>, como la del *Epistolario*, que se llevó a cabo con la Fundación Universitaria Española.<sup>2</sup>

---

<sup>1</sup> **Menéndez Pelayo, Marcelino.** *Obras completas.* - Santander : Consejo Superior de Investigaciones Científicas, 1942-1962 (Aldus). - 65 v.

<sup>2</sup> **Menéndez Pelayo, Marcelino.** *Epistolario* / edición al cuidado de Manuel Revuelta Sañudo ; advertencia preliminar del Excmo. Sr. D. Pedro Sáinz Rodríguez. - Madrid : Fundación Universitaria Española, 1982-1991. - 23 v, - (*Documentos históricos* ; 12)

Ambas instituciones han colaborado desinteresadamente en esta, por así decirlo, segunda edición de sus respectivas publicaciones y es de justicia reconocerlo aquí.

Además, se ha contado, desde el primer momento, con la participación de la Fundación Histórica Tavera,<sup>3</sup> que desarrolla un amplio programa de edición de fuentes documentales y bibliográficas en formato electrónico, relacionadas con la cultura española e iberoamericana, y la de la Asociación de Hispanismo Filosófico<sup>4</sup>, con quien ya hubo una primera colaboración consistente en la publicación de las Actas de las III Jornadas<sup>5</sup> de esta asociación. Ambas instituciones, junto con la Biblioteca de Menéndez Pelayo, han diseñado un proyecto más amplio y ambicioso denominado *Polígrafos*, que se expone en el prólogo de las mencionadas *Actas*. La Obra Social de la Caja de Ahorros de Santander y Cantabria ha corrido muy generosamente con la mayor parte de los gastos.

#### EL PROYECTO *POLÍGRAFOS ESPAÑOLES E IBEROAMERICANOS*

El proyecto *Polígrafos Españoles e Iberoamericanos* aspira, como objetivo general, a hacer accesible internacionalmente el pensamiento de figuras españolas, iberoamericanas o de otras naciones, que sean autores de obra amplia en extensión, diversa en temática y de trascendencia en el desarrollo de proyectos culturales. A estos destacados personajes se alude bajo la denominación genérica de "Polígrafos".

Para ello, se considera la edición de un amplísimo conjunto de autores de interés fundamental tanto para la cultura española, europea e iberoamericana. La lista abarca desde las

<sup>3</sup> <http://www.tavera.com>.

<sup>4</sup> <http://www.ffil.uam.es/AHF>

<sup>5</sup> *Estudios sobre historia del pensamiento español : (actas de las III Jornadas de Hispanismo Filosófico) I* Antonio Jiménez García (editor). - Santander : Asociación de Hispanismo Filosófico, Sociedad Menéndez Pelayo, Fundación Histórica Tavera, 1998. - IX, 406 p. ; 24 cm. - {*Estudios de literatura y pensamiento hispánicos*}. - ISBN 84-86993-20-2.

obras de los grandes polígrafos españoles de carácter universal (entre ellos, Ramón Llull, Luis' Vives, Francisco Suárez); los españoles, desde la época romana hasta el siglo XX, sin olvidar a los polígrafos portugueses, brasileños o iberoamericanos.

La vinculación de este proyecto con la Biblioteca de Menéndez Pelayo, tanto en su aspecto científico como operativo, es evidente. Por una parte, Menéndez Pelayo constituye el primero de los polígrafos españoles —y uno de los más grandes a escala universal— y, por otra, la Biblioteca Menéndez Pelayo es depositaria de un extenso y notable fondo de obras de "polígrafos" españoles e iberoamericanos. Además, no es desdeñable el papel impulsor de estos proyectos para la introducción y aplicación de la más avanzada tecnología de impresión digital, que tan amplia repercusión tendrá en el siglo XXI, en este caso en la Comunidad de Cantabria.

El objetivo específico del proyecto será la preparación conjunta, por las instituciones participantes, de una serie de volúmenes digitales (cederrón) en los que de forma completa o antológica se incluyan obras representativas de Polígrafos, para ponerlas a disposición de instituciones académicas y culturales de España, Portugal e Iberoamérica, constituyendo de este modo un extenso patrimonio bibliográfico, al que en la actualidad difícilmente puede accederse por tratarse de obras agotadas o de elevado coste en el mercado, lo que, en general, dificulta no sólo su adquisición sino incluso su consulta. Como objetivo último e ideal, y para completar satisfactoriamente la repercusión exterior del proyecto, se considera la necesidad de hacer accesible en Internet el conjunto de información digitalizada resultante. La inclusión en esta vía esencial de la comunicación de un "corpus" tan amplio de la cultura española e iberoamericana redundará, sin duda, en un aporte sustancial a la presencia en Internet de nuestra cultura, uno de los retos más importantes que en la actualidad afronta la Comunidad Cultural Iberoamericana.

## LAS OBRAS COMPLETAS DE MENÉNDEZ PELAYO

En este marco más general es donde se indicarla la edición de las obras completas de Marcelino Menéndez Pelayo. En este trabajo se pretende únicamente dar a conocer los aspectos técnicos del proyecto, pues se trata de una de las primeras ocasiones en las que se lleva a cabo de una forma tan masiva la edición electrónica de un autor español (no olvidemos que las *Obras completas* abarcan 65 volúmenes y el *Epistolario*, 23). A diferencia del proyecto de digitalización del *Boletín de la Biblioteca de Menéndez Pelayo*, que se describe más adelante, de forma complementaria, los escritos de Menéndez Pelayo se ha realizado en modo texto, no en imagen, es decir utilizando códigos ASCII. No deja de ser llamativo que los 88 volúmenes de apretada tipografía quepan perfectamente en un solo cederrón. En concreto 30.214 páginas corresponden a las *Obras completas* y 12.829 al *Epistolario*. Más tarde todo, todo ese conjunto, se ha estructurado con el lenguaje HTML.

Esta edición electrónica cuenta con el apoyo financiero de la Caja de Ahorros de Santander y Cantabria. De hecho, el correspondiente convenio se firmó el 7 de abril de 1999, coincidiendo con la celebración de las IV Jornadas de la Asociación de Hispanismo Filosófico.

El proceso ha consistido en el escaneado de la totalidad de las páginas, aplicándose posteriormente un programa de reconocimiento óptico de caracteres para transformar la imagen en códigos ASCII. Indudablemente este proceso ha necesitado de un trabajo cuidadoso de revisión de la transcripción del texto original, así como de presentación del mismo. Inicialmente, y para facilitar la rápida recuperación del texto, se ha utilizado como lenguaje de marcación, el nivel más sencillo y de más amplio uso en Internet, HTML, v.3, que en sucesivas ediciones avanzará a un etiquetado en XML. Esta labor de ecdótica electrónica permitirá añadir una mayor potencia y pertinencia a la recuperación de la información. Específicamente dentro de XML, el seguimien-

to de las directrices TEI<sup>6</sup> para la codificación de textos electrónicos posibilitará esta tarea y permitirá la recuperación de la información en un nivel mucho más extenso que el actual; de hecho, ya está anunciado que las nuevas versiones de Internet Explorer o Netscape lleven añadida la funcionalidad de recuperación y visualización de documentos XML, sin necesidades de *plugins* especiales.

Se han añadido, además, campos específicos de búsqueda para extraer toda la potencialidad de recuperación posible. En suma, la totalidad de la obra de Marcelino Menéndez Pelayo puede ser interrogada, seleccionada y visualizada según los títulos de la obra, los resúmenes, los capítulos y epígrafes, las notas a pie de página, y lo que es más importante por cualquier palabra del texto. Las consultas según estos campos y con el auxilio de los operadores booleanos pueden aplicarse tanto a la totalidad de las obras como a un subconjunto. Todo ello abre multitud de posibilidades de investigación y evitará muchos trabajos tediosos de búsqueda y cotejo. Este es, sin lugar a dudas, el motivo por el cual está teniendo tanta repercusión la edición de textos electrónicos<sup>7</sup> fundamentales para cada una de las especialidades del conocimiento científico.

Efectivamente, la necesidad de índices para el manejo en profundidad de estas obras ha obligado en muchos casos a trabajos complementarios de indización, que se han llevado lar-

<sup>6</sup> *TEI Guidelines for Electronic Text Encoding and Interchange (P3)* / Made available by the Electronic Text Center at the University of Virginia. <http://etext.virginia.edu/TEI.html>.

<sup>7</sup> Uno de los sitios web más amigables es el del proyecto de la Universidad de Oxford Oxford Text Archive, <http://firth.natcorp.ox.ac.uk/ota/public/index.html>.

Otros sitios web clásicos en este tema son el Center for the Electronic Texts in the Humanities (CETH) que ofrece un directorio de centros de textos electrónicos. <http://www.ceth.rutgers.edu/informationServices/ectrdir.html>.

En el ámbito hispano ver el Proyecto Filosofía en Español. <http://www.uniovi.es/~filespe>.

gas horas de esfuerzo de muchos investigadores. Simplemente, la combinación de la edición electrónica, los lenguajes de marcación apropiados y los programas de recuperación basados en ellos eliminan la necesidad de estos índices puesto que permiten tanto la búsqueda a texto completo como la consulta de palabras, conceptos y nombres contextualizada según las marcas lógicas establecidas. En el caso de las obras de Menéndez Pelayo, su significación viene dada no sólo por la importancia de su obra sino también por la enorme cantidad de referencias bibliográficas e informaciones referencia-les que aporta. Estas referencias, convenientemente etiquetadas, como se hará en las siguientes fases del proyecto, permitirán navegar entre los distintos títulos, en un modo de biblioteca virtual, tal y como el mismo maestro habría deseado. Como ejemplo, *La ciencia española*<sup>8</sup> es un magnífico inicio de un inventario general de la ciencia y el pensamiento español. Cabe decir que se abren unas insospechadas posibilidades para el estudio intrínseco de la obra de un autor a través del marcado exhaustivo de sus textos pues se pueden estudiar sus influencias, *clusters* de citas, evolución de vocabulario, persistencia de temas, etc., etc., desde una perspectiva más rigurosa.<sup>9</sup>

Hay que añadir, además, que la misma línea establecida para las *Obras completas* de Menéndez Pelayo de difusión de las fuentes documentales y bibliográficas de la Biblioteca de Menéndez Pelayo se está siguiendo tanto en la digitalización del *Boletín* como en la digitalización de los *Papeles* y codifi-

<sup>8</sup> **Menéndez Pelayo, Marcelino.** *La ciencia española* / edición ordenada y dirigida por Don. Miguel Artigas. - Madrid : Victoriano Suárez, 1933. -2v.'; 25 cm.

<sup>9</sup> Sobre este aspecto podrá consultarse el artículo de Xavier Agenjo, *La edición digital de los textos filosóficos españoles*, que se ha presentado a las Cuartas Jornadas de la Asociación de Hispanismo Filosófico. *Actas de las IV Jornadas de Hispanismo Filosófico*. - Santander : Sociedad Menéndez Pelayo : Asociación de Hispanismo Filosófico, 2000. - En prensa

cación en EAD<sup>10</sup> de su *Inventario*.<sup>11</sup> No cabe duda que este proyecto guarda relación con la Red de Centros y Archivo Virtual de la Cultura Española Contemporánea (1868-1936) y que la metodología llevada a cabo en la digitalización de las obras de Menéndez Pelayo tiene importantes similitudes con la que se empleará en la mencionada iniciativa. Incluso, y lo que es más importante, ambos proyectos guardan una identidad absoluta en lo que se refiere a normativa. Necesariamente habrán de interrelacionarse, siempre que el alcance de los escritores representados en cualquiera de ellos termine por coincidir; tal vez el caso de Rafael Altamira pueda dar pie a una primera colaboración.

Como se ha dicho antes, de forma complementaria, pero con distinta metodología, se ha llevado a cabo, de forma simultánea, el proyecto de la edición digital del *Boletín de la Biblioteca de Menéndez Pelayo*, que se publicará por las mismas fechas que *Menéndez Pelayo digital*.

El *Boletín de la Biblioteca de Menéndez Pelayo*<sup>12</sup> se comenzó a editar en 1919<sup>13</sup> y desde entonces se ha venido publicando ininterrumpidamente, excepto durante el período de la Guerra Civil, tras la que incluso se promovió un cambio de cabecera que, afortunadamente, no llegó a cuajar, reanudándose su marcha regular. Hasta este año han aparecido 75

---

10 *Las ISDB(ER) y la EAD : un Caso Práctico 64th IFLA I* Martínez-Conde, María Luisa y Agenjo Bullón, Xavier // 64<sup>th</sup> IFLA General Conference, August 16 - August 21, 1998 : Conference Programme and Proceedings. <http://www.ifla.org/IV/ifla64/066-74s.htm>.

11 <http://www.unican.es/bibmp/inv-int.htm>.

12 *Boletín de la Biblioteca de Menéndez Pelayo*. - En.-feb. 1919- . - Santander : Sociedad Menéndez Pelayo, 1919- . - v. ; 24 cm. - Bimestral ; anual. - Disponible también en microficha. - ISSN 0006-1646 = *Boletín de la Biblioteca de Menéndez Pelayo*.

13 Para más detalles sobre la historia del *Boletín* puede consultarse el artículo de Manuel Revuelta Sañudo *La Biblioteca de Menéndez Pelayo // Boletín de la Biblioteca de Menéndez Pelayo*. - Santander : Sociedad Menéndez Pelayo. - ISSN 0006-1646. - LV (1979) 297-300.



tomos y se encuentra en prensa el 76, correspondiente a 1999. La revista cuenta con una amplia difusión en su campo especializado de actuación, como lo prueba el elevado número de citas que suelen recibir los artículos que allí se publican, en las numerosas publicaciones especializadas en la cultura y literatura españolas. Otra prueba de su repercusión es la altísima tasa de canje con otras publicaciones similares y que alcanza en 1999 el centenar de títulos.

Existe una fuerte implantación histórica de esta revista en muchas bibliotecas. En efecto, si se consulta el Catálogo Colectivo Español de Publicaciones Periódicas, a fecha 21 de junio de 1999 aparecen hasta un total de 77 colecciones, aunque no debe olvidarse que la actualización del Catálogo Colectivo no es aún completa.<sup>14</sup> Justamente y como resultado de la consulta a este Catálogo en línea se pudo comprobar que era bastante considerable el número de instituciones que, por múltiples razones, poseían colecciones incompletas del *Boletín*, lagunas que se conocen técnicamente según la norma ISO 10324 como *gap*.

Esta situación era ya de sobra conocida por los bibliotecarios de la Biblioteca de Menéndez Pelayo que habían tenido que contestar en muchas ocasiones de forma negativa a las solicitudes de algunas bibliotecas, que requerían números

<sup>14</sup> La última aportación sobre el estado de la cuestión puede consultarse en *Nueva fase del Catálogo Colectivo de Publicaciones Periódicas : el intercambio de información inform.atiza.da* / Francisca Hernández Carrascal, Xavier Agenjo Bullón // *Archivos, bibliotecas, centros de documentación y museos en el Estado de las autonomías : actas del VI Congreso Nacional de ANABAD : fMurcia, 26, 27 y 28 de junio de 1996*. - Murcia : ANABAD-Murcia ; Consejería de Cultura y Educación, 1997. - p. 331-338. Desde entonces, se ha puesto en funcionamiento la base de datos del Catálogo Colectivo Español de Publicaciones Periódicas formado por 70.677 registros bibliográficos, 278.635 registros de fondos pertenecientes a 1.140 bibliotecas (<http://www.bne.es/esp/ccpp.htm>). Esperamos que cuando este texto se publique se haya finalizado la actualización de los registros procedentes de las bibliotecas del CSIC.

sueltos para completar sus colecciones, o de nuevos suscriptores que deseaban adquirir los números previos. La falta de ejemplares ha impedido en la mayoría de los casos atender estas peticiones.

Por esta razón hace ya algunos años que se tomó la decisión de microfilmear la colección completa en 103 microfichas,<sup>15</sup> solución muy bien recibida sobre todo en universidades americanas, bien en sus bibliotecas o en sus departamentos de español. Sin embargo, las limitaciones de las microfichas son en la actualidad perfectamente subsanables y para ello se tomó la decisión de digitalizar la colección completa, de forma tal que se satisficieran todas las demandas posibles. Esta decisión se tomó en el marco más amplio de la nueva política de difusión de textos electrónicos adoptada por la Biblioteca de Menéndez Pelayo. De hecho, y como se acaba de relatar desde 1997 ya se estaba trabajando con la Fundación Histórica Tavera<sup>16</sup> en el primer paso de un ambicioso proyecto que recibe el nombre de *Polígrafos españoles* y sobre el cual se dio una primera información en el VII Congreso de ANABAD celebrado en Toledo.<sup>17</sup>

Por lo tanto, de forma simultánea al proyecto *Polígrafos españoles* se abordó el proyecto de digitalización del *Boletín de la Biblioteca de Menéndez Pelayo*. Para ello se ha contado, además de la aportación de la Fundación Histórica Tavera, con el apoyo del Ayuntamiento de Santander. En cualquier caso, y meramente a efecto ilustrativo para quien desee emprender proyectos similares, conviene advertir que el presupuesto no ha superado apenas los dos millones de pesetas.

<sup>15</sup> *Boletín de la Biblioteca de Menéndez Pelayo [Microforma]*. - En.-feb. 1919-. - Santander : Sociedad Menéndez Pelayo, 1919-. - microfichas ; 11 x 15 cm.

<sup>16</sup> Para más información sobre esta Fundación puede consultarse su página web <http://www.tavera.com>.

<sup>17</sup> *La biblioteca virtual de Menéndez Pelayo. El proyecto Polígrafos I* Xavier Agenjo, Francisca Hernández // *La información y el derecho de los ciudadanos : VII Congreso Nacional de ANABAD*. [En prensa].

Los ejecutores materiales de la operación han sido los informáticos de la empresa Digibis, ligada a la Fundación Histórica Tavera, a la que se le debe ya una impresionante lista de publicaciones en soporte digital, tal vez no suficientemente reconocida.<sup>18</sup>

Desde el punto de vista técnico se optó por una digitalización en modo imagen con una resolución de 400 ppp en formato TIFF, con compresión Grupo IV, tanto para la digitalización de los textos como para las láminas que acompañan en ocasiones a los artículos. El total de la información de las cerca de 30.000 páginas está al macenado en tres cederrones y está previsto que en la próxima actualización, que se publicará el año 2000, incluyendo la digitalización de tres nuevos números del *Boletín*, que el soporte elegido sea el DVD.

Está claro que el acceso a la información contenida en unas páginas digitalizadas en modo imagen, y no en modo texto, depende de la indexación de los contenidos; de otro modo no se obtienen ventajas superiores a la paginación tradicional. Afortunadamente, el *Boletín* tenía este trabajo hecho, ya que Andrés del Rey y Rosa Fernández publicaron unos magníficos índices<sup>19</sup> que ampliaban y mejoraban considerablemente el que en 1938 elaboró Tomás Maza y el que en 1959 reunió Ignacio Aguilera, tercer director de la Biblioteca de Menéndez Pelayo. Estos índices son sumamente amplios y contienen varios subíndices, el de artículos, donde se reflejan los autores, títulos, número y páginas del *Boletín* en el que aparecieron publicados; el índice de obras reseñadas en el que aparecen los autores de las obras reseñadas, sus títulos, editores literarios, traductores, prologuistas, etc., así como

---

<sup>18</sup> Información sobre la empresa puede obtenerse en <http://www.digibis.com>. Una lista completa de sus publicaciones está en <http://www.digibis.com/digibis/productos.htm>.

<sup>19</sup> *Índices del Boletín de la Biblioteca de Menéndez Pelayo (1919-1995)* / Andrés del Rey Sayagués, Rosa Fernández Lera // *Boletín de la Biblioteca de Menéndez Pelayo*. - Santander : Sociedad Menéndez Pelayo. - ISSN 0006-1646. - LXXII (1996) 269-416.

los autores de las reseñas; el índice de la Crónica, en el que se referencian los actos, homenajes, reuniones y noticias internas de la Biblioteca; y por último, y como complemento a los otros, el índice general de materias permite un acceso temático al contenido completo del *Boletín*.

También están consultables en línea en la página web de la Biblioteca de Menéndez Pelayo <<http://www.unican.es/bibmp/boletinl.htm>> y han probado su eficacia a través del elevado número de consultas que indican los marcadores estadísticos de esta página y, sobre todo, a través de las solicitudes de documentos que han generado. Aunque el *índice* apareció en el tomo de 1996, lógicamente se ha actualizado con las nuevas entradas correspondientes a los años 1997 y 1998.

Está claro que la digitalización de una publicación periódica en modo imagen se convierte en un magnífico instrumento de acceso a la información sólo si va provisto de unos índices completos. Cuanto mayor sea el grado de exhaustividad de los referidos índices mayor será la eficacia en la búsqueda de los contenidos. En este caso contamos, como se ha dicho, con unos índices considerablemente completos, lo cual no excluye que en una fase posterior del proyecto se lleve a cabo un reconocimiento óptico de caracteres de las imágenes digitalizadas, que permitirán la búsqueda sobre la totalidad del texto del Boletín. No debe olvidarse que para obtener buenos resultados en la aplicación de OCR sobre imágenes digitales, éstas debe alcanzar un umbral mínimo de resolución que se sitúa en torno a los 400 ppp. Por esta razón se optó desde el primer momento por un nivel tan alto de resolución.

Como se ha dicho al principio, la premisa básica que ha empujado a la digitalización del *Boletín* ha sido la difusión de su contenido a los ciudadanos y, en términos generales, la mejora del acceso a las publicaciones conservadas en la Biblioteca. Una duda frecuente en torno a la digitalización es la conveniencia o no de poner en línea el contenido comple-

to de una o varias, en ocasiones muchas, publicaciones. Esta duda está motivada por los problemas de derechos de autor, de costes, de control del número de accesos, relacionados con la decisión, pero en ocasiones también está producida por un planteamiento muy conservador en cuanto a la difusión de los contenidos. En el caso del *Boletín* no existen problemas de derechos de autor sobre la publicación, ya que la editora, la Sociedad Menéndez Pelayo, continua siendo la editora de la versión electrónica de esta publicación y su proyecto prevé actualizar esta versión, con posterioridad a la tradicional edición en papel. Es decir, se pretende la coexistencia de ambas versiones. En cuanto a los costes, ya se ha mencionado el presupuesto con el que cuenta el proyecto, y se estima una reducción drástica de los esfuerzos de personal en el proceso de respuesta a las solicitudes de números atrasados. Con todo ello, la influencia en los costes no puede ser considerada sólo en términos directos, sino a la luz de la mejora del acceso a la información que representa, y los beneficios indirectos que conlleva.

Es frecuente que el tema del acceso en línea a los contenidos de publicaciones digitalizadas lleve aparejado una discusión acerca de la conveniencia del acceso libre a las mismas, transportando a Internet la misma polémica acceso libre/acceso restringido que se debate en algunas bibliotecas con fondos especiales. A nuestro modo de ver esta discusión no es aplicable en absoluto a Internet y supone una aplicación poco justa de los presupuestos del Estado con los que funcionamos. Otra visión, también restrictiva aunque mucho más vistosa, es la de alimentar los sitios *web* con algunas de las publicaciones de mayor interés, supuestamente, a modo de escaparate virtual de una institución. Con ello, se corre el peligro de que una vez compuesto el escaparate y demostrado el prestigio de una institución con algunas imágenes, se abandonen los proyectos e, ipso facto, el objetivo de facilitar a los ciudadanos el acceso a la información.

La resolución de esta discusión debe basarse en criterios técnicos, para lo cual nos permitimos aportar una pequeña sugerencia metodológica. ¿Hasta qué punto será necesario poner en línea el contenido completo de la digitalización del *Boletín* o mantenerlo exclusivamente en soportes DVD?. Continuando con la línea de facilitar el acceso a la información, en lo que no hay ninguna duda es en las enormes ventajas que tiene que los índices sean accesibles en Internet, pero, tal vez, la velocidad de transmisión de la información en la Internet de 1999 suponga un freno a su consulta. En cambio, el almacenamiento de la información en un servidor no parece presentar muchas dificultades dado el abaratamiento de la memoria. Sin duda, se podrá profundizar técnicamente en esta cuestión analizándola en base a las *vatios* entre el volumen de información y costes en euros por número de consultas y recuperaciones de información a través de la línea. Parece claro que tendrá que existir una relación directa entre el índice de consulta de la información y la *vatio* de disponibilidad. Estos aspectos de orden técnico, así como sus perspectivas de futuro, se analizan a continuación.

Durante muchos años la práctica archivística, bibliotecónica y museística, ha seguido caminos aparentemente diferenciados, fortaleciéndose una separación profesional basada en la incompatibilidad de los tratamientos documentales, cualitativamente distintos, y en la impermeabilidad de los conocimientos de estas ciencias, derivada también de las diferencias en los tratamientos. Sin embargo, estas diferencias están siendo eliminadas por Internet y por la renovación profesional, tecnológica y técnica que ésta indudablemente conlleva. Todo el desarrollo de las tecnologías de la información, redes de comunicaciones, los lenguajes de marcación (o marcado, pues aún no está consolidada la traducción de *mark-up languages*), la apertura e interconexión de sistemas, etc., ha provocado en 5 años, por no remontarse más allá en el tiempo, el estallido de nuevas posibilidades de transmisión, búsqueda y

recuperación de documentos. Y esta vez no solo de documentos secundarios, o terciarios, sino fundamentalmente de documentos primarios.

Paralelamente, junto con las capacidades de la red, se han empezado a ver satisfechas las necesidades de los ciudadanos de acceder a todo tipo de documentos. Investigadores y usuarios, la sociedad en general, demandan más documentos en la red, mayor capacidad de acceso, más facilidad de manejo. Así lo consideran los profesionales e instituciones más vanguardistas, que están superando, por la fuerza de los hechos, las históricas resistencias de muchos profesionales de la archivística, de la biblioteconomía, de la documentación y de la museística a utilizar una normativa común y una misma metodología que facilite la consulta de documentos primarios a los ciudadanos. Por un lado, tenemos las iniciativas perfiladas en el V Programa Marco<sup>20</sup> de la Unión Europea, en el que justamente se pretende fomentar esta armonización de criterios con el objetivo de alcanzar una sociedad de la información fácilmente manejable. Por otro lado, existen numerosos proyectos de acceso a documentos primarios, animados por grandes redes de bibliotecas y archivos<sup>21</sup> y las grandes asociaciones internacionales como IFLA,<sup>22</sup> CÍA<sup>23</sup> e ICOM<sup>24</sup> han visto también la necesidad de esta colaboración.

Existe pues, una profesión, varias profesiones, centradas en el acceso a los documentos, tres vías de acceso al conocimiento. ¿Existe una cuarta? ¿Quién cabe en ella? ¿Qué líneas

---

<sup>20</sup> Para más información sobre este programa pueden consultarse las siguientes direcciones: y <http://www.bne.es/esp/general.htm> y <http://www.cordis.lu/fp5/src/over.html>.

<sup>21</sup> En esta actividad está siendo muy activa el Research Library Group (RLG), su programa para el acceso a documentos primarios puede consultarse en: <http://www.rlg.org/primary/index.html>.

<sup>22</sup> <http://www.ifla.org>.

<sup>23</sup> <http://www.archives.ca/ica>.

<sup>24</sup> ICOM - The International Council of Museums <http://www.icom.org>.

de trabajo han de asumir las tres tradicionales profesiones de archiveros, bibliotecarios y museólogos? Las tres trabajan con documentos primarios, las tres elaboran documentos secundarios y terciarios. Es más, el núcleo fundamental de su actividad es la elaboración de éstos documentos, conforme a las normas, técnicas y prácticas que les impone la profesión y, mucho más a menudo, en función de unos recursos muy limitados, en los que la falta de comunicación disminuye las capacidades de estas profesiones.

Inventarios, registros y catálogos, bases de datos, ficheros manuales, listados, cuadernos, etc., en los que se trata de una información que, cruzada, podría dar unos resultados impresionantes. Se puede imaginar los resultados de una búsqueda sobre Francisco de Goya en la que recuperaríamos material de archivo de todo tipo, objetos conservados en museos, obra gráfica, pinturas, frescos, reproducciones, películas, videos, etc. Una visión comprehensiva de Francisco de Goya. ¿Es esto realmente posible? Y si la pregunta fuera, por ejemplo, los recursos culturales de una comunidad (municipio, provincia, comunidad autónoma, nación, etc.), y si existiera la posibilidad de interconectar sistemas utilizando Z39.50 y localizar todo lo existente de Goya en todo el mundo...

¿Tan difícil resultará adoptar estructuras de información similares de modo tal que se permita el intercambio de información y el intercambio de descripciones entre objetos heterogéneos?. Las descripciones de objetos heterogéneos pueden presentar uniformidades que reduzcan las diferencias a algo homogéneo. Estas homogeneidades son básicamente títulos, autores, entidades, materias, es decir puntos de acceso, aunque existen muchas más. En este sentido, la biblioteconomía ha avanzado de manera considerable en la función de control a los puntos de acceso, produciendo un entramado de normativa dedicada a la elaboración de registros de autoridad, a su codificación, y a su presentación e intercambio.



## CONDICIONANTES BÁSICOS DEL SISTEMA

Existen una serie de condicionantes sobre las posibles soluciones técnicas a adoptar hasta conseguir un sistema de tales características, es decir, un acceso global y simultáneo a archivos, bibliotecas y museos. Estos condicionantes vienen definidos básicamente por los conjuntos documentales en sí mismos y por la normativa existente. Hay que mencionar que en el caso de Menéndez Pelayo, su obra y su biblioteca, sus papeles y su Casa-Museo, constituyen un magnífico campo para esta visión holística de acceso a la información.

### A. LOS CONJUNTOS DOCUMENTALES

Son la base material de un sistema global, al que condicionan por las necesidades de tratamiento que exigen y por las normas y prácticas aplicables para la elaboración de descripciones automatizadas. Este punto está íntimamente relacionado con el análisis de la normativa existente y no puede desligarse de él.

De un modo genérico la base documental de un sistema de acceso global a archivos, bibliotecas y museos, estará formada por:

#### *Material archivístico:*

Archivos históricos, administrativos, personales y familiares, con toda la variedad de documentos y soportes que los conformen, sin olvidar la propia documentación electrónica. Un análisis de los papeles del archivo puede consultarse en su catálogo-inventario.<sup>25</sup>

---

<sup>25</sup> *Catálogo-inventario de los manuscritos y papeles de la Biblioteca de Menéndez Pelayo (2ª parte)* / Manuel Revuelta Sañudo, Rosa Fernández Lera y Andrés del Rey Sayagués. - Santander : Sociedad Menéndez Pelayo, 1994. - ISBN 84-605-0694-0.

*Material bibliográfico:*

Manuscritos, monografías, publicaciones periódicas, material gráfico. Evidentemente pueden existir recursos electrónicos fruto de la digitalización de diversos materiales de carácter audiovisual, pues el principio de estos soportes y medios de expresión, vienen a coincidir con las últimas dos décadas de la vida de don Marcelino.

*Material museológico:*

Obras de arte o artefactos, o por mejor decir, sus representaciones digitales: imagen, sonido, movimiento, texto, etc., así como los objetos digitales complejos formados por la combinación de varios tipos. No abundan las descripciones de las piezas de la Casa-Museo, aunque en la primera parte del *Catálogo de manuscritos* de Miguel Artigas, a lo largo del *Boletín*<sup>26</sup> y en la *Guía de la Biblioteca* de Manuel Revuelta,<sup>27</sup> pueden encontrarse interesantes noticias.

**B. OBJETOS DEL SISTEMA Y ESTRUCTURA BÁSICA**

La definición de los objetos sobre los que trabajará el sistema se rige a su vez por la normativa para la elaboración de las descripciones informatizadas archivísticas, bibliográficas y museísticas. Se trata de describir conjuntos documentales heterogéneos, a los que, por tanto, deberá aplicárseles procedimientos de catalogación heterogéneos, pero que, dependiendo de las normas utilizadas para ello, pueden conformar registros de un mayor nivel de homogeneidad. La homogeneidad de las descripciones tendrá en la capa superior de la sintaxis de los registros, o documentos electrónicos descrip-

---

<sup>26</sup> Cfr. nota 19.

<sup>27</sup> Cfr. nota 13.

tivos, una enorme influencia en la capacidad de búsqueda y recuperación del sistema.

Existe la opción de crear una base de datos para el material de archivo, otra para el bibliográfico y otra para el museístico, ya que constituyen descripciones heterogéneas, o bien, una sola base de datos con las descripciones homogéneas de ambos materiales.

La primera opción está claramente limitada por la propia normativa archivística y museística, que no ofrece aún ni un sistema de codificación, ni un formato de comunicación para el intercambio de información. No debe confundirse la necesidad de un formato común de codificación y comunicación para el intercambio de información en los procesos de búsqueda y recuperación, con el intercambio destinado a procesos de, por ejemplo, catalogación cooperativos. De hecho, en la historia de la normalización en bibliotecas ha tenido mayor influencia la necesidad de difundir información a través de catálogos en fichas primero y bases de datos después, que el objetivo de intercambio en procesos cooperativos, que ha sido alcanzado posteriormente. Es frecuente que se achaque el retraso en la normalización de archivos y museos, y en concreto la falta de un formato de comunicación de información, al hecho de trabajar con documentos únicos cuyas descripciones no son intercambiables por definición. Sin embargo, esta explicación, que podría ser válida para el intercambio de descripciones, no ofrece ninguna razón que aclare la falta de descripciones informatizadas con fines meramente informativos.

Esto supone que el desarrollo o aplicación de algún sistema de gestión automatizada de archivos sólo puede llegar a ser normalizado en la descripción de unidades archivísticas, o de registros de autoridad de archivo, si incluye las normas *ISAD(G)* e *ISAAR (CPF)*. En el caso de los museos, la norma de referencia básica son las *International guidelines for Museum Object Information* del CIDOC. En ambos casos, como se ha dicho, no existe aún un formato de codificación sobre el que basar sistemas de búsqueda y recuperación con una sintaxis común.

## NORMATIVA

Para el entorno archivístico existe, como es sabido, la norma *ISAAR(CPF)*.<sup>28</sup> Esta norma, aprobada en 1995, establece unas reglas mínimas para la creación de registros de autoridad de archivos. El interés fundamental de esta norma es el control de los puntos de acceso principales de los instrumentos de descripción archivísticos, es decir, las entidades, personas y familias creadoras de archivos.

Esta norma tienen un matiz importante que la diferencia de las normas aplicables a los registros de autoridad bibliográficos y es que su finalidad es, no solo normalizar la forma de los nombres de los creadores de archivos, sino la de describir todos aquellos atributos de éstos que son necesarios para apreciar el contexto de creación de los documentos. Por tanto, tiene una vinculación contextual con el registro archivístico mucho mayor que la que existe entre los registros de autoridad y los registros bibliográficos.

Sin embargo, como se indica en su *Introducción*, punto 11, esta norma sólo aporta una parte de los requisitos necesarios para el intercambio de información de autoridades de archivo. Las normas y convenciones para la normalización de puntos de acceso deben ser las establecidas a nivel nacional (*AACR2*, *APPM* (*Archives, Personal Papers, and Manuscripts*), *Reglas de Catalogación*, *AFNOR NF Z 44-061*, etc.) Además, para que sirva realmente para el intercambio de información todavía tiene que adoptar o desarrollar un formato de comunicación, o bien, utilizar alguno ya existente. Por supuesto que la apuesta más segura es la ISO 2709 y la ISO 23950.

<sup>28</sup> *ISAAR(CPF) International Standard Archival Authority Record for Corporate Bodies, Persons and Families* : final ICA approved versión / prepared by the Ad Hoc Commission on Descriptive Standards, París, France, 15-20 November 1995. - Ottawa : The Secretariat of the ICA Ad Hoc Commission on Descriptive Standards, 1996.

Por otra parte, toda la información necesaria para crear un registro de autoridad de archivo puede ser incluida sobradamente en un registro de autoridad bibliográfico, por lo que la recomendación sería que el trabajo de control de autoridades incorpore, cuando sea necesario, la información adicional relevante requerida por la norma ISAAR (CPF) que pueda no estar presente en los registros de autoridad bibliográficos. No parece razonable suponer que esa recomendación vaya a encontrar ningún tipo de resistencia. En este sentido, un buen aprovechamiento de los recursos aconseja la utilización de *Autoridades de la Biblioteca Nacional en cd-rom* como base de datos de registros de autoridad, que en su última actualización supera la cifra de 250.000 registros y que cubre las necesidades expuestas por las ISAAR(CPF). No es necesario insistir que el caso del control de puntos de acceso en museos podría beneficiarse del uso de los mismos recursos, y permitiría dedicar el esfuerzo de construcción de bases de datos de autoridades a otros destinos.

#### COMPARACIÓN DE NORMAS:

- La normativa bibliotecaria es mucho más amplia ya que abarca la normalización de los puntos de acceso de nombres de persona, entidades, congresos, títulos, materias y nombres geográficos, lo que cubre todas las necesidades de bibliotecas, archivísticos y museos. La normativa archivística y museística para puntos de acceso está en un estadio mínimo de desarrollo y aplicación, aunque puede experimentar un gran avance utilizando las bases de datos de autoridad ya existentes.
- Dispone de un formato de comunicación para el intercambio de registros (MARC,<sup>29</sup> ISO 2709) tanto para los

---

cfr. MARC21.

procesos de información, como para los procesos de creación e intercambio de registros de autoridad.

- Existen grandes bases de datos de registros de autoridad que permiten la captura de registros (*Autoridades de la Biblioteca Nacional en CD-ROM, Library of Congress Authority Names, Les Notices d'Autorité de BN-OPALE sur CD-ROM, etc.*) a partir de cederrón o en línea (Z39.50).
- La normativa archivística para registros de autoridad no cubre todos los tipos de puntos de acceso posible (solo personas, familias y entidades), y por tanto, su capacidad para la recuperación o para la navegación es limitada.
- Los ficheros de autoridad automatizados que puedan cumplir la norma ISAAR(CPF) no están en disposición de constituir una base de acceso a un sistema global de búsqueda y recuperación.

A la vista de lo que precede sería muy recomendable una revisión de ambos conjuntos de normas con un espíritu amplio: la conjunción de ambas es algo inevitable. Y el proceso se va a experimentar de forma natural en la Biblioteca de Menéndez Pelayo Sin duda, tiene mucho más que ver con la impermeabilidad de los profesionales de ambos mundos que con la impermeabilidad real de la práctica, y sobre todo, con los deseos de investigadores, usuarios y de la sociedad en general. Tampoco es desdeñable el esfuerzo que ha de realizar cualquier comisión normalizadora, y mucho más si es internacional, que pretenda establecer mínimos comunes en un mundo no interconectado. ¿Es realmente aceptable la regla no escrita que dice que no puede haber interconexión entre aquellos que tratan con documentos únicos y, por el contrario, es posible normalizar el mundo de los que trabajan con copias de un mismo documento.

De los conjuntos descriptivos básicos se extraerá una serie de elementos comunes que permitan el acceso global y simultáneo a las bases de datos, independientemente de su origen, materiales componentes y descripciones. En este sentido, existen dos aproximaciones a la resolución de estos problemas, basadas ambas en la convergencia y complementariedad de normas. Consisten en la ampliación del marco Z39.50, por un lado, y en la aplicación de lenguajes de marcación, concretamente de las definiciones de tipo de documentos SGML específicas para archivos (EAD), bibliotecas (MARC DTD) y museos (MUS-EAD).

De hecho, desde hace unos años se han venido definiendo perfiles específicos<sup>30</sup> dentro de la norma Z39.50 aplicables tanto al acceso, búsqueda y recuperación en colecciones digitales, de bibliotecas y archivos, así como de museos y recursos electrónicos relacionados con el patrimonio cultural (lo que incluye arte, arquitectura, historia de la cultura, historia natural, etc.). El perfil Z39.50 para acceso a colecciones digitales es el marco al que pertenecen otros subconjuntos, como el Perfil *CIMI* (*Consortium for Interchange of Museum Information Profile*), y pretende solventar los problemas de acceso entre colecciones digitales heterogéneas. Este perfil define como categorías de objetos digitales los basados en el lenguaje, en imágenes o en sonido y movimiento; y en las categorías de descripciones asociadas trabaja tanto con registros bibliográficos, como con registros de archivos o cabeceras de documentos HTML y páginas web. Siguiendo el mismo sistema, el perfil CIMI, a su vez, utiliza los atributos definidos en Bib-1, según el consenso esta-

<sup>30</sup> Consúltase *Z39.50 Profile for Access to Digital Collections : Draft Seven : (Final Draft for Review)* May 3, 1996. - [Washington] : Library of Congress, 1996. - <http://lcweb.loc.gov/z3950/agency/profiles/coHections.himl>.

*The CIMI Profile : Z39.50 Application Profile for Cultural Heritage Information : Release 1.0 March 1, 1998* / Prepared by Consortium for the Computer Interchange of Museum Information (CIMI) CIMI Z39.50 Working Group. - <http://www.cimi.org/downloads/ProfileFinalMar98/cimi-profilel.htm>.

blecido por la Conferencia de Bibliotecas Nacionales Europeas (CENL),<sup>31</sup> para que los clientes Z -basados en CIMI puedan acceder a bases de datos bibliográficas y para que los servidores CIMI puedan así mismo ser accesibles por clientes Z de orientación bibliográfica.<sup>32</sup>

Dentro del mundo bibliotecario existen dos normas que regulan la creación de registros de autoridad informatizados: las Reglas de Catalogación y el formato IBERMARC; dentro del mundo archivístico se está extendiendo la norma ISAAR (CPF) para archivos. ¿Existe conciliación? Tal vez la solución sean los lenguajes de marcación, en el que el tándem SGML/XML tiene mucho que ofrecer en los próximos años. En este marco, es especialmente significativa la rápida extensión de la codificación EAD,<sup>33</sup> definición de tipo de documento lanzada y mantenida conjuntamente por la *Library of Congress* y la *Society of American Archivists*, que tanto interés ha despertado también en el mundo museológico, aunque no sin ciertas reticencias. Debe recordarse que la EAD codifica en SGML los denominados instrumentos de descripción (*finding aids*), entre los cuales se cuentan los catálogos e inventarios de archivos y

---

31 *The CENL Z39.50 profile : Versión 1.2, 14 December 1998*. <http://linnea.helsinki.fi/z3950/cenlprofile.html>.

32 En cualquier caso, el estado de la cuestión sobre la normalización de museos, un buen paso adelante, pero solo eso, puede consultarse en *Normalización documental de museos: elementos para una aplicación informática de gestión museográfica I* Andrés Carretero Pérez [et al.]. - Madrid : Ministerio de Educación y Cultura, 1998. En las páginas 549 a 560 se presenta una completa bibliografía relacionada con el objeto de este artículo. No obstante, es posible hacer fuertes salvedades en este documento que no puede considerarse en modo alguno como definitivo.

33 Encoding Archival Description. <http://lcweb.loc.gov/ead>. Una descripción de los proyectos EAD que se están llevando a cabo en todo el mundo está en <http://www.loc.gov/ead/eadsites.html>.

*La descripción archivística codificada : un formato internacional para la descripción archivística* / Ana Teresa García Martínez, Rosario Jiménez Vela y Consuelo de las Mercedes Martín Vega // *X Jornadas Bibliotecarios de Andalucía, Jerez de la Frontera, 1998*. - Málaga : Asociación Andaluza de Bibliotecarios, 1998. - p. 411-430.



museos.<sup>34</sup> Esta línea de convergencia tiene también como característica el aprovechamiento de las bases de datos bibliográficas y de autoridades existentes, para extenderlas en una doble dirección. Por un lado, son la base del control de los contenidos codificados en elementos comunes de acceso (<persname>, <entname>, <geoname>, <date>, <titleproper>, <edition>, etc., entre otros muchos), y por otro ofrece una norma común para la codificación de descripciones de bibliotecas, archivos y museos en profundidad (SGML) y para su difusión en Internet a través de XML.

En España podría citarse el proyecto ARCCa ('Acceso público a los Recursos Culturales de Cantabria')<sup>35</sup> que se elabora en la Comunidad Autónoma de Cantabria y que justamente, en este momento, se enfrenta con ese tipo de problemas. Tiene como objetivo principal la difusión y preservación del patrimonio cultural bibliográfico, museístico y archivístico de esa región, con el fin de facilitar a los ciudadanos el acceso a los recursos culturales de la Comunidad. Las dos grandes líneas de actuación de ARCCA serán la automatización e interconexión de bibliotecas, archivos y museos de la Comunidad y la digitalización de los materiales en ellos depositados. Por otro lado, la iniciativa cántabra encontró eco en proyectos similares de la Unión Europea desarrollán-

---

34 El estado de la cuestión en la normalización internacional referida a museos puede seguirse en el sitio del The International Committee for Documentation of the International Council of Museums (ICOM-CIDOC) <http://www.cidoc.icom.org/#standardse>.

Ver también MUS-EAD Museums & the Encoded Archival Description, iniciativa del Museum Computer Network, <http://world.std.com/~mcn/standards/mus-ead.html>.

35 Ya se puede consultar los primeros estadios de este proyecto en la dirección <http://www.arcca-cantabria.org>. Para un resumen de este proyecto puede consultarse *Proyecto ARCCA: Acceso Público a los Recursos Culturales de Cantabria* ; Xavier Agenjo, José María Gutiérrez // *Correo bibliotecario : Boletín informativo de la Subdirección General de Coordinación Bibliotecaria*. - ISSN 1136-2111. - N° 13 (marzo 1997). - <http://www.bcl.jcyl.es/correo/Correol3/Mundol3.html>.

dose entonces el denominado proyecto *GeniusLoci* que, aunque *retenido* por la *Comisión* en la convocatoria anterior, se ha preferido trasladar, a instancias de la propia Comisión a la presente convocatoria, descrita en el párrafo Iº de esta comunicación, pues, en último término, su objetivo no es otro que el de desarrollar una herramienta informática que permita a través de una único interface que el ciudadano ejercite su derecho a la información patrimonial y cultural independientemente que esta se encuentre conservada en un archivo una biblioteca o un museo.

Buena prueba de que la sociedad demanda este tipo de servicios de información es el proyecto basado en las mismas premisas expuestas en esta comunicación presentado el 18 de enero de 1999 por D. José María Aznar, Presidente del Gobierno, D. Emilio Botín, presidente del día anterior fusionado Banco Santander Central Hispano-Americano (BSCH), la entonces Ministra de Cultura y hoy Presidenta del Senado D<sup>a</sup> Esperanza Aguirre y el Director de la Residencia de Estudiantes D. José García Velasco, todo lo cual le proporcionó una inusitada atención de los medios de comunicación, tan infrecuente en las presentaciones de proyectos culturales, la *Red de Centros y Archivo Virtual de la Cultura Española Contemporánea (1868-1936)*. Buena parte de aquella filosofía y de lo que se ha venido exponiendo se ha trasladado al proyecto COVAX (contemporary virtual archives in *xml*) en el que la Residencia de Estudiantes, como coordinadora, y la Biblioteca de Menéndez Pelayo, entre otras instituciones europeas, han formado un consorcio que ha logrado finalmente la correspondiente subvención dentro del Plan Marco IST (*Information Society Technologies*)<sup>36</sup>.

XAVIER AGENJO BULLÓN  
Biblioteca de Menéndez Pelayo  
FRANCISCA HERNÁNDEZ CARRASCAL  
Institución Libre de Enseñanza